Supplemental Appendix

A. Details of Estimation Method

A.1 Numerical solution

We solve the model for value functions and policy functions with the collocation method in Miranda and Fackler (2004).

A.1.1 Problem

The recursive problem during one generation is

$$
\begin{aligned}
V(a,r,w,t) &= \max_{c} \mathbf{1}\{t<T\}\left\{u(c)+\beta V(a',r,w,t+1)\right\}+\mathbf{1}\{t=T\}\left\{u(c)+e(a')\right\} \\
s.t. & \\
a' &= (1+r)(a-c)+w \\
c &\leq a \\
c &\geq 0
\end{aligned}
$$

The problem can be written as

$$
\begin{aligned}
V_1(a,r,w) &= \max_{c\in[0,a]} u(c)+\beta V_2((1+r)(a-c)+w,r,w) \\
V_2(a,r,w) &= \max_{c\in[0,a]} u(c)+\beta V_3((1+r)(a-c)+w,r,w) \\
&\vdots \\
V_{T-1}(a,r,w) &= \max_{c\in[0,a]} u(c)+\beta V_T((1+r)(a-c)+w,r,w) \\
V_T(a,r,w) &= \max_{c\in[0,a]} u(c)+e((1+r)(a-c)+w)
\end{aligned}
$$

The parameters are: $\{\beta,T,u(c),e(a)\}$. Set $T=6$ for simplicity and we can decrease β to account for the longer length of periods.

A.1.2 Collocation

The state space is $s=(a,z)$. $z=(r,w)$ is the exogenous state which has the transition matrix $P=P_r\otimes P_w$. The state space for z is discrete and so is enumerated $k=1,\ldots,K$, where $K=N_r\times N_w$. Let

$s = (s_1, s_2)$ and the choice variable $x = c$. The choice is consumption $x \in B(s)$, where

$$B(s) = [0, a]$$

Re-writing this as a system of six value functions

$$\begin{aligned} V_1(s) &= \max_{x \in B(s)} F_1(s,x) + \beta V_2([(1+r)(s_1 - x) + w, s_2]) \\ &\vdots \\ V_T(s) &= \max_{x \in B(s)} F_2(s,x) \end{aligned}$$

This is the system we will solve.

Approximation: Take $V_1, \ldots, V_T$ and approximate them on J collocation nodes $s_1, \ldots, s_J$ with a spline with J coefficients $c^1 = (c^1_1, \ldots, c^1_J)$, $c^2, \ldots, c^T$ and linear basis ϕ_j.

$$\begin{aligned} V_1(s_i) &= \sum_{j=1}^{J} c^1_j \phi_j(s_i) \\ &\vdots \\ V_T(s_i) &= \sum_{j=1}^{J} c^T_j \phi_j(s_i) \end{aligned}$$

Let $c = (c^1, \ldots, c^T)$ and let $v_1(c^1) = [V_1(s_1), \ldots, V_1(s_J)]'$ and $v_2(c^2), \ldots, v_T(c^T)$ similarly defined for a given c. With $v(c) = [v_1(c^1)', \ldots, v_J(c^J)']'$ then

$$\begin{aligned} v_1(s) &= \Phi c^1 \\ &\vdots \\ v_T(s) &= \Phi c^T \end{aligned}$$

this is the *collocation equation*.

Substituting the interpolants into the value functions

$$\begin{aligned} \sum_{j=1}^{J} c^1_j \phi_j(s_i) &= \max_{x \in B(s_i)} F_1(s_i, x) + \beta \sum_{j=1}^{J} c^2_j \phi_j([(1+r)(s_{i,1} - x) + w, s_{i,2}]) \\ \sum_{j=1}^{J} c^2_j \phi_j(s_i) &= \max_{x \in B(s_i)} F_1(s_i, x) + \beta \sum_{j=1}^{J} c^3_j \phi_j([(1+r)(s_{i,1} - x) + w, s_{i,2}]) \end{aligned}$$

$$\vdots$$

$$\sum_{j=1}^{J} c_j^T \phi_j(s_i) = \max_{x \in B(s_i)} F_2(s_i, x)$$

The stacked system of value functions is

$$\begin{aligned}
\Phi(s)c^1 &= F_1(s, x(s)) + \beta\Phi([(1+r)(s_1 - x(s)) + w, s_2])c^2 =: v_1(c^2) \\
\Phi(s)c^2 &= F_1(s, x(s)) + \beta\Phi([(1+r)(s_1 - x(s)) + w, s_2])c^3 =: v_2(c^3) \\
&\vdots \\
\Phi(s)c^T &= F_2(s, x(s))
\end{aligned}$$

The zero system would be $\tilde{\Phi}(s)c - v(c) = 0$, where $\tilde{\Phi}$ is a block diagonal matrix of $\Phi's$.

A.2 Estimation procedure

The estimation procedure is described as below in two steps, adapted from Guvenen (2016). The global stage is a multi-start algorithm where candidate parameter vectors are uniform Sobol (quasi-random) points. We typically take about 10,000 initial Sobol points for pre-testing and select the best 200 points (i.e., ranked by objective value) for the multiple restart procedure. The local minimization stage is performed with the Nelder-Mead's downhill simplex algorithm (which is slow but performs well on non-linear objectives). Within one evaluation, we draw 100,000 individuals randomly and simulate their entire wealth process initiated with zero wealtha and the lowest earnings profile.

A.3 Additional results

We report the full mobility matrix here for the baseline:

$$\hat{T}_{36} = \begin{bmatrix}
0.274 & 0.247 & 0.238 & 0.152 & 0.041 & 0.043 & 0.005 \\
0.246 & 0.263 & 0.254 & 0.145 & 0.048 & 0.036 & 0.009 \\
0.252 & 0.233 & 0.269 & 0.147 & 0.058 & 0.033 & 0.008 \\
0.238 & 0.254 & 0.239 & 0.158 & 0.066 & 0.039 & 0.009 \\
0.201 & 0.266 & 0.262 & 0.164 & 0.047 & 0.049 & 0.013 \\
0.232 & 0.250 & 0.240 & 0.152 & 0.040 & 0.041 & 0.044 \\
0.207 & 0.303 & 0.172 & 0.082 & 0.030 & 0.084 & 0.122
\end{bmatrix}$$

The full matrices for all other cases are available upon request.

B. Input Data Sources

B.1 Labor income levels

The labor income data we use is adapted from the PSID data cleaned by Heathcote et al. (2010), specifically Sample C in their labeling. We only keep those aged between 25-60 inclusively. Then we construct the age-dependent decile values in the following order: this order corresponds to several implicit assumptions, the most important of which is that we allow people to move across bins during their life cycle.

1. for each age calculate the decile values of earnings;

2. for each age bin of six years, calculate the average decile earnings across these six years.

The above order maintains the distributional ranking of model agents across the life cycle.

B.2 Intergenerational labor income transitions

Chetty et al. (2014) provide a 100 by 100 transition matrix linking parental family income and child's income in their online data and tables, with each cell corresponding to share of each percentile of the income distribution.[1] The main sample they use is the Statistics of Income (SOI) annual cross-sections from 1980 to 1982 cohorts for children, and the authors link children to their parents using population tax records spanning 1996-2012 for parent family income. We collapse this big matrix into a 10 by 10 transition matrix, with each cell corresponding to share of each decile of the income distribution. Note that this matrix captures intergenerational transition in income.

Online table 2 of Chetty et al. (2014) also provide the average income levels for both parent and child. However, they are an average income around a particular age (29-30) for both parent and child rather than an average life cycle income. We would like our income profiles to capture the hump-shaped life cycle feature, thus calculate our own as explained in the last sub-section.

C. Output Data Sources

C.1 Wealth distributional moments

The wealth distributional moments are taken from Díaz-Giménez et al. (2011). Their calculations are

[1] See `http://equality-of-opportunity.org/images/online_data_tables.xls`, online table 1.

more cleaned and serve as an official report. Many papers have used their numbers, e.g. in Kindermann and Krueger (2015). Other estimates are very close.

C.2 Intergenerational Wealth mobility moments

There are three papers, to the best of our knowledge, that estimate a transition matrix for wealth mobility. Kennickell and Starr-McCluer (1997) and Klevmarken et al. (2003) are both estimated using panel data, i.e. not necessarily transition across generations. The former paper used SCF panel and the latter used PSID panel. Charles and Hurst (2003) are a transition matrix for generations in particular. Please note the difference, though I try to argue they yield similar estimates.

Kennickell and Starr-McCluer (1997)calculate the six-year transition matrix from 1983 to 1989 for quartiles and top percentile ranges, and their results are quite similar to Klevmarken et al. (2003). The seven states are: bottom 25, 25-49, 50-74, 75-89, 90-94, top 2-5, top 1, respectively. Their estimates are (from Table 7),

$$T_{KS,6} = \begin{bmatrix} 0.672 & 0.246 & 0.063 & 0.018 & 0.001 & 0.000 & 0.000 \\ 0.246 & 0.495 & 0.190 & 0.042 & 0.019 & 0.007 & 0.000 \\ 0.066 & 0.192 & 0.480 & 0.208 & 0.037 & 0.016 & 0.000 \\ 0.021 & 0.082 & 0.329 & 0.418 & 0.113 & 0.036 & 0.002 \\ 0.011 & 0.071 & 0.212 & 0.301 & 0.225 & 0.177 & 0.004 \\ 0.000 & 0.028 & 0.164 & 0.104 & 0.180 & 0.430 & 0.094 \\ 0.000 & 0.031 & 0.024 & 0.061 & 0.045 & 0.247 & 0.593 \end{bmatrix}$$

When raised to the power of 6 (i.e. 36-year transition matrix), we have

$$T_{KS,36} = \begin{bmatrix} 0.316 & 0.278 & 0.222 & 0.118 & 0.037 & 0.024 & 0.005 \\ 0.276 & 0.263 & 0.240 & 0.137 & 0.044 & 0.031 & 0.009 \\ 0.224 & 0.242 & 0.263 & 0.163 & 0.054 & 0.042 & 0.012 \\ 0.196 & 0.229 & 0.274 & 0.176 & 0.061 & 0.051 & 0.013 \\ 0.179 & 0.219 & 0.275 & 0.181 & 0.066 & 0.061 & 0.020 \\ 0.150 & 0.198 & 0.271 & 0.185 & 0.074 & 0.082 & 0.040 \\ 0.112 & 0.166 & 0.252 & 0.182 & 0.085 & 0.121 & 0.083 \end{bmatrix}$$

We see that the 36-year transition matrix does not necessarily reach the stationary distribution.

Klevmarken et al. (2003)calculate the five-year transition matrix from 1994 to 1999 for quintiles using

PSID data. Note the states for the Markov chain are different. Their estimates are (from Table 6),

$$T_{KLS,5} = \begin{bmatrix} 0.583 & 0.273 & 0.099 & 0.031 & 0.015 \\ 0.267 & 0.435 & 0.223 & 0.058 & 0.016 \\ 0.087 & 0.208 & 0.419 & 0.232 & 0.055 \\ 0.048 & 0.079 & 0.193 & 0.481 & 0.200 \\ 0.014 & 0.022 & 0.051 & 0.200 & 0.713 \end{bmatrix}$$

One potential issue with the above transition matrix is that it does not necessarily capture the *inter-generational* transmission in wealth. For that argument, let us look at the alternative transition matrix estimated by Charles and Hurst (2003).[2] There are two transition matrices in Table 5. If we only adjust the logs of parental and child wealth for age, the matrix is:

$$T_{CH,gen} = \begin{bmatrix} 0.23 & 0.21 & 0.18 & 0.21 & 0.17 \\ 0.25 & 0.17 & 0.19 & 0.21 & 0.19 \\ 0.20 & 0.25 & 0.20 & 0.20 & 0.15 \\ 0.15 & 0.17 & 0.21 & 0.21 & 0.25 \\ 0.17 & 0.20 & 0.22 & 0.17 & 0.24 \end{bmatrix}$$

If we adjust logs of parental and child wealth for "age, income, and portfolio choice," the corresponding matrix is:

$$T_{CH,gen,adj} = \begin{bmatrix} 0.36 & 0.29 & 0.16 & 0.12 & 0.07 \\ 0.26 & 0.24 & 0.24 & 0.15 & 0.12 \\ 0.16 & 0.21 & 0.25 & 0.24 & 0.15 \\ 0.15 & 0.13 & 0.20 & 0.26 & 0.26 \\ 0.11 & 0.16 & 0.14 & 0.24 & 0.36 \end{bmatrix}$$

with each cell corresponding to a quintile-to-quintile transition probability. Again note the differences in the states of the Markov chain.

[2]Sample selection: Their sample consists of all PSID parent-child pairs in which (a) the parents were in the survey in 1984–89 and were alive in 1989, (b) the child was in the survey in 1999, (c) the head of the parent family was not retired and was between the ages of 25 and 65 in 1984, (d) the child was between ages 25 and 65 in 1999, and (e) both the child and the parent had positive wealth when measured. There were 1,491 such parent-child pairs.

References

Charles, Kerwin Kofi and Erik Hurst, "The Correlation of Wealth across Generations," *Journal of Political Economy*, December 2003, *111* (6), 1155–1182.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Díaz-Giménez, Javier, Andrew Glover, and José-Víctor Ríos-Rull, "Facts on the distributions of earnings, income, and wealth in the United States: 2007 update," *Quarterly Review*, 2011.

Guvenen, Fatih, *Quantitative Economics with Heterogeneity: An A-to-Z Guidebook*, Princeton University Press, 2016.

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante, "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006," *Review of Economic Dynamics*, January 2010, *13* (1), 15–51.

Kennickell, Arthur B and Martha Starr-McCluer, "Household Saving and Portfolio Change: Evidence from the 1983-89 SCF Panel," *Review of Income and Wealth*, December 1997, *43* (4), 381–99.

Kindermann, Fabian and Dirk Krueger, "High Marginal Tax Rates on the Top 1%? Lessons from a Life Cycle Model with Idiosyncratic Income Risk," 2015.

Klevmarken, Anders, Joseph P. Lupton, and Frank P. Stafford, "Wealth Dynamics in the 1980s and 1990s: Sweden and the United States," *Journal of Human Resources*, 2003, *38* (2).

Miranda, Mario J. and Paul L. Fackler, *Applied Computational Economics and Finance*, Vol. 1 of *MIT Press Books*, The MIT Press, June 2004.